MSA

July 25, 2020

AUCKLAND HOUSE PRICES ANALYSIS

[1]:
```
import json
import sys
sys.path.append('/home/nbuser/library/')

import pandas as pd
import requests
```

[2]:
```
def get_pop(Longitude,Latitude):



    url = 'https://koordinates.com/services/query/v1/vector.json'
    params ={
     'key': '15e07883fca44c0195654266ee7ebe17',
     'layer': 104612,
     'x': Longitude,
     'y': Latitude
    }
    results = requests.get(url=url, params=params)

    if results.status_code != 200:
        return pd.Series({'Population':results.status_code})

    pop = results.
 ↪json()['vectorQuery']['layers']['104612']['features'][0]['properties']['C18_CURPop']


    return pd.Series({'Population':pop})
```

[3]:
```
df = pd.read_csv('Dataset%20for%20Assignment.csv')
```

[4]:
```
df['Population']= df.apply(lambda row:␣
 ↪get_pop(row['Longitude'],row['Latitude']),axis=1)
```

[5]:
```
index = pd.read_csv('otago730395.csv')
```

[6]:
```
index = index[['SA12018_code','NZDep2018']]
```

[7]:
```
df = df.merge(index,left_on='SA1' ,right_on='SA12018_code')
```

[8]:
```
df.to_csv('Completed_dataset.csv',index=False)
```

[9]:
```
dataset = pd.read_csv('Completed_dataset.csv')
```

[10]:
```
dataset = dataset.drop('SA12018_code', axis=1)
```

[11]:
```
dataset.isnull().values.any()
```

[11]:
```
True
```

[12]:
```
dataset.isnull().sum()
```

[12]:
```
Bedrooms       0
Bathrooms      2
Address        0
Land area      0
CV             0
Latitude       0
Longitude      0
SA1            0
0-19 years     0
20-29 years    0
30-39 years    0
40-49 years    0
50-59 years    0
60+ years      0
Suburbs        1
Population     0
NZDep2018      0
dtype: int64
```

[13]:
```
dataset = dataset.fillna({'Bathrooms': 0})
dataset = dataset.dropna(axis=0,how='any')
dataset.isnull().sum()
```

[13]:
```
Bedrooms       0
Bathrooms      0
Address        0
Land area      0
CV             0
Latitude       0
Longitude      0
SA1            0
0-19 years     0
20-29 years    0
30-39 years    0
40-49 years    0
50-59 years    0
60+ years      0
Suburbs        0
Population     0
NZDep2018      0
```

```
dtype: int64
```

[14]:
```
dataset.dtypes
```

[14]:
```
Bedrooms         int64
Bathrooms      float64
Address         object
Land area       object
CV               int64
Latitude       float64
Longitude      float64
SA1              int64
0-19 years       int64
20-29 years      int64
30-39 years      int64
40-49 years      int64
50-59 years      int64
60+ years        int64
Suburbs         object
Population       int64
NZDep2018      float64
dtype: object
```

[15]:
```
dataset['Land area'] = dataset['Land area'].str.extract('(\d+)').astype('float')
```

[16]:
```
dataset.to_csv('House_Price.csv',index=False)
```

[ ]:
```

```

[ ]:
```

```